

Quality control trials of prothrombin time:

An assessment of the performance in serial studies

L. POLLER, JEAN M. THOMSON, AND K. F. YEE

From the National (UK) Reference Laboratory for Anticoagulant Reagents and Control, Haematology Department, Withington Hospital, Manchester M20 8LR, UK

SUMMARY A series of collaborative exercises on the one-stage prothrombin time test involving hospitals in Britain and overseas was performed between 1972 and 1977. The British Comparative Thromboplastin (BCT) and the lyophilised test plasmas were issued from the National (UK) Reference Laboratory for Anticoagulant Reagents and Control. Participants were asked to test the plasma samples with the BCT using the recommended technique. Variability of performance was assessed by the 'index of reliability' based on the plasma variance and error variance within each exercise. The results show that hospitals have attained higher precision in the later trials.

A series of quality control trials of the prothrombin time technique was conducted by the National (UK) Reference Laboratory for Anticoagulant Reagents and Control to assess hospitals' performance with the British Comparative Thromboplastin (BCT) using the recommended technique for the one-stage prothrombin time test. In a study of earlier trials performed in 1972-73 (Leck *et al.*, 1974), hospitals' results varied considerably within each trial, and only a very slight indication of improvement with successive trials was found.

Results from further proficiency studies extending over a five-year period are now available. The aim of this survey is to reappraise the variability of the hospitals' performance with the BCT and the recommended technique for the prothrombin time in the programme of collaborative proficiency exercises.

Method of study

The National (UK) Reference Laboratory for Anticoagulant Reagents and Control has conducted national and international collaborative studies in blood coagulation at regular intervals since 1972. These have included nine similar designed quality control of performance studies for the prothrombin time test (serial numbers 5, 7, 8, 9, 11, 14, 17, 18, and 19) which were conducted from 1972 to 1977. In each of these trials, involving hospitals in Britain and overseas, supplies of BCT, together with three

lyophilised test plasmas, the necessary diluents, and calcium chloride, were sent to participants. The lyophilised plasmas represented different degrees of coagulation defect. Instructions for reconstituting and testing the plasmas and a report form were also provided. On each occasion participants reported the test results with the three plasmas as clotting times together with their own local normal control values.

The object of these exercises has been to monitor individual hospitals' performance of the prothrombin time test. After every trial a print-out was sent to hospitals showing the result as a percentage deviation for each plasma and the average deviation for all three plasmas. The histogram of the average deviation of all participants was also included so that an individual hospital could compare its performance with the overall distribution of results. Results were expressed as prothrombin times and prothrombin ratios.

A further aim was to determine whether there had been any improvement in reliability of results. An ideal solution would be to find a single index to assess the overall performance for every trial. We have found that plasmas with severe coagulation defects tend to have higher standard deviations. Therefore, the standard deviations of results of tests were not directly comparable because different plasmas were used in each exercise. The inclusion of the coefficient of variation (CV), the other commonly used parameter which measures precision in our report of each trial, does not overcome this difficulty. Therefore, a formula, termed variance stabilising transformation, which allows comparison of SD of

Received for publication 27 September 1978

individual plasmas, was used. This is derived as follows:

$$\frac{100}{\sqrt{\text{prothrombin time}}}$$

Using an analysis of variance, the variation of the prothrombin time measurement can be broken down into two components: the plasma variance (arising from the use of different plasmas) and the error variance, that is, the variation due to technical and random fluctuation (see Addendum). If all the hospitals' results were the same, there would be no error variance, and the variability of the results would be due to plasma differences only (plasmas tested were not the same within each trial). A ratio of the plasma variance to the total variance can be used to assess the role played by the differences in plasmas in the different trials in the production of the total variation. From this ratio an index termed 'the index of reliability' can be used to assess the consistency of the hospitals' results and is defined as $R = \sqrt{(\text{plasma variance}/\text{total variance})}$ (Guilford, 1965). R lies between zero and unity and the higher the value R, the closer the agreement between hospitals in the trial.

Results

Table 1 gives the plasma variance and error variance for each trial from which the R values have been determined. Hospitals showed less variation in the later prothrombin time proficiency trials. In the Figure, the index of reliability (R) is plotted for each trial. The index value increased at trial 9 and remained relatively unchanged onwards. To see whether the increase is significant, a linear regression line has been calculated. The R values show a significant increasing trend ($P < 0.01$).

Table 1 Variance components, index of reliability (R), and number of hospitals in quality control trials

Trial No.	No.	Plasma variance	Error variance	R
5	199	12	1.6	0.936
7	252	32	3.4	0.952
8	267	8	1.4	0.925
9	344	41	2.6	0.970
11	344	40	1.9	0.977
14	330	38	1.9	0.976
17	340	29	1.6	0.974
18	266	43	2.8	0.968
19	359	36	1.2	0.984

Using the CUSUM (cumulative sum technique) curve, further confirmation was obtained that a significant change occurred at trial 9 ($P < 0.005$).

The means and standard deviations of the results

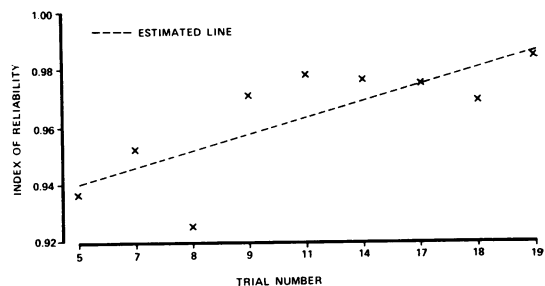


Figure Indices of reliability (R) of the series of prothrombin time trials.

Table 2a Mean (SD) of prothrombin times of test plasmas used in series of prothrombin time quality control trials

Trial No.	Plasma mean (SD)		
5	23.1 (3.0)	34.5 (4.2)	51.4 (8.4)
7	12.2 (1.4)	24.8 (3.9)	32.2 (7.9)
8	22.1 (3.0)	25.6 (3.4)	40.6 (6.5)
9	12.0 (1.6)	27.2 (3.0)	36.2 (7.7)
11	13.7 (2.1)	30.1 (5.4)	45.0 (6.4)
14	12.9 (1.1)	29.5 (4.3)	38.1 (6.0)
17	14.2 (1.7)	33.3 (4.2)	33.8 (4.2)
18	12.0 (2.0)	29.5 (6.5)	35.7 (5.8)
19	12.9 (1.7)	29.5 (2.9)	35.1 (3.2)

Table 2b Mean (SD) of transformed data corresponding to Table 2a

Trial No.	Plasma mean (SD)		
5	21 (1.4)	17 (1.0)	14 (1.3)
7	29 (1.5)	20 (1.4)	18 (2.4)
8	21 (1.1)	20 (1.2)	16 (1.2)
9	29 (1.4)	19 (1.2)	17 (2.1)
11	27 (1.4)	18 (1.3)	15 (1.4)
14	28 (1.1)	19 (1.5)	16 (1.4)
17	27 (1.3)	17 (1.2)	17 (1.2)
18	29 (1.5)	19 (2.1)	17 (1.3)
19	28 (1.3)	18 (1.0)	17 (1.0)

for each plasma used in the trial are shown in Table 2a. The three plasmas showed different degrees of coagulation defect in each trial. The means and SD of the transformed data are also given (Table 2b). The difference between SD became negligible when the data were transformed in the way indicated (Table 2b).

Discussion

We are encouraged by the improvement of performance shown in the serial prothrombin time trials. The purpose of quality control of performance studies, as well as to monitor the performance, should be to improve the standard of reliability. It

is, therefore, gratifying that this second objective appears to have been achieved in this progressive series of exercises.

In a previous study (Leck *et al.*, 1974) we showed that there was considerable variation in the measurement of the prothrombin time among the hospitals in the earlier trials; this may have been due to the lack of familiarity with the BCT and the recommended technique for the one-stage prothrombin time test. If this was the case, after long-term participation in a series of quality control trials, the hospitals might improve their performance, resulting in greater precision of measurement. From the Figure the increasing trend of R (index of reliability) indicates that the performance in the later trials is better than in the earlier ones. Also the absence of change in index values in the later trials may suggest that a plateau of performance has been reached.

We have learned from previous experience that some plasmas give a larger variation than others. This makes comparison of different trials difficult. To overcome the problem arising from the fact that the test plasmas used in the serial trials were always different, several transformations have been employed in this analysis which gave similar results showing significant increasing trends of the index of reliability. The transformation using the formula based on $100/\sqrt{\text{prothrombin time}}$ is given in detail in this report because it showed the least dependency ($r = -0.04$) on the means and SD of the plasmas.

The closer agreement shown in the later international prothrombin time proficiency trials demonstrates that the BCT is being successfully employed by the many participant centres in Britain and overseas and that the series of quality control studies has resulted in a gratifying improvement of performance.

We express our gratitude to Dr M. K. Palmer, Christie Hospital and Holt Radium Institute, for advice on the presentation of the statistical analysis,

and to Dr P. J. Laycock, University of Manchester Institute of Science and Technology, for an independent assessment of the statistical methods used.

References

- Guilford, J. P. (1965). Reliability of measurement. In *Fundamental Statistics in Psychology and Education*, 4th edition, pp. 438-466. McGraw-Hill, New York.
- Leck, I., Gowland, E., and Poller, L. (1974). The variability of measurements of the prothrombin time ratio in the National Quality Control Trials: a follow-up study. *British Journal of Haematology*, **28**, 601-612.

Requests for reprints to: Dr L. Poller, National (UK) Reference Laboratory for Anticoagulant Reagents and Control, Withington Hospital, Manchester M20 8LR, UK.

Addendum

The statistical model was derived as follows:

$$Y = \mu + P + H + I + E$$

where

Y = measurement,	variance σ_Y^2
P = plasma effect,	variance σ_P^2
H = hospital effect,	variance σ_H^2
I = interaction term,	variance σ_{PH}^2
E = residual noise,	variance σ_R^2

The σ_{PH}^2 and σ_R^2 terms are confounded, hence we are led to examine the value for $\sigma_H^2 + \sigma_{PH}^2 + \sigma_R^2 = \sigma_e^2$ say the error variance which measures the precision of the hospitals within each trial. However, the sequence of the error variances cannot be used to compare the hospital's performance because the variance stabilising transformation of individual plasmas also stabilises the error variances. The total variance of a measurement is $\sigma_Y^2 = \sigma_P^2 + \sigma_e^2$. Here we assume the plasma effect is a random variable.

$R^2 = \sigma_P^2/(\sigma_P^2 + \sigma_e^2)$ is the interclass correlation coefficient, that is, the ordinary correlation coefficient between any two hospitals' measurements of the same plasma.